

Autonomous object modeling based on affordances in a dynamic environment

Simon L. Gay¹ and Salima Hassas²

¹ Université Lyon 1, Villeurbanne, France
`simon.gay@liris.cnrs.fr`

² Université Lyon 1, Villeurbanne, France
`salima.hassas@univ-lyon1.fr`

Abstract

We present an architecture for self-motivated agents to generate behaviors in a dynamic environment according to its possibilities of interactions. Some interactions have predefined valences that specify inborn behavioral preferences. Over time, the agent learns to recognize affordances in its surrounding environment under the form of structures called signatures of interactions. The agent keeps track of enacted interactions in a spatial memory to generate a completed context in which it can use signatures to recognize and localize distant possibilities of interactions, and generates behaviors that satisfy its motivation principles.

Keywords: Affordances, Autonomous learning, Body schema, Developmental learning, Constructivism learning, Self-motivation, Spatial awareness.

1 Introduction

In this paper, we address the problem of the construction, interpretation and exploitation of a short-term representation of a dynamic environment by an artificial agent that initially ignore elements that compose its environment. Our work is based on a model proposed by Georgeon and Aha [3], called Radical Interactionism, in which actions and perceptions are considered as inseparables and kept embedded within data structures called *interactions*. Specifically, interactions are used to model Piaget's notion of sensorimotor scheme [8].

In previous works [2][1], we proposed and tested mechanisms that allow an artificial agent to construct and exploit such a structure to generate behaviors adapted to the environmental context of the agent. These mechanisms are inspired from biology: most natural organisms have brain structures that maintains some geometrical correspondences with the animal's surrounding space, such as tectum[6] (or colliculus in mammals), or sensorimotor cortices[5]. These mechanisms are however limited to a static environment. To overcome this limitation, we propose to adapt these mechanisms to consider movements of elements of the environment. This paper focuses on the adaptation of mechanisms that learn to define objects that compose the environment, and mechanisms that generate behaviors.

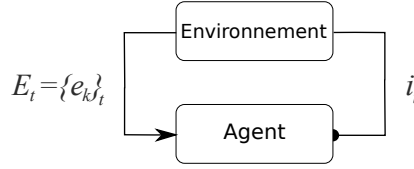


Figure 1: Model of the Parallel Radical Interactionism [2]. A decision cycle begins with an *intended interaction* i_t of the agent, which then experiments a set of *enacted interactions* E_t

2 Formalization of Parallel Radical Interactionism (PRI)

The PRI model[2] begins with a set I of *primitive interactions*. Each primitive interaction i is attributed a *valence* v_i that defines the agent’s behavioral preferences. This principle defines a form of intrinsic motivation [7], called *Interactional Motivation*. A PRI decision cycle, illustrated in Figure 1, begins with an *intended interaction* i_t from the agent. At the end of the cycle, the agent is informed of interactions that were actually enacted, called *Interactional Context* E_t .

We thus consider that the agent receives additional sensory *stimuli* when an interaction is enacted. These *stimuli* cannot be considered without the interaction that produce them. We thus propose to construct new interactions by associating an interaction and a *stimuli*. We call *secondary interaction* i such an interaction, *associated interaction* of i the interaction that composes i , and *primary interaction* an interaction that is not based on another one.

In a preliminary work¹, we have implemented an agent equipped with a sequential RI algorithm that can capture sequential properties of mobile elements. We thus propose that the agent can consider sequences of two enacted interactions. Indeed, a sequence of interactions can carry information about the relative movement of an object detected by these interactions.

3 The Spatial Memory System

We have proposed a structure, called *Space Memory System* (SMS) [2][1], that completes the interactional context E_t and helps the agent to generate behaviors adapted to the environment.

The SMS is based on structures that can characterize elements of the environment, called *Signatures of Interaction*. The principle is based on the assumption that the result of the enaction of an interaction depends on a limited context of elements of the environment, considered as the “objects” that affords this interaction. This definition of objects relates to the concept of affordance proposed by Gibson[4]: an *object* is a specific spatial configuration of elements in the environment that affords an interaction, and does not require *a priori* knowledge. A RI agent cannot directly observe these objects, but can “experience” them through interactions. The aim of the signature mechanism is thus to define for each interaction i a set of interactions $\{j\}$ for which the enaction can characterize the presence of the object that affords i . This mechanism is similar to the mechanism proposed by Uğur et al. [9]. However, the use of interactions rather than actions and perceptions makes it possible to learn spatial properties of space, as interactions carry information about the movements of the agent in space. Our mechanism is adapted from [2], and extended to sequences of interactions. The mechanism is formalized as a *certitude function* c that gives the certitude that an interaction or sequence i can be successfully enacted in a context E with an absolute certitude of success when $c(i, E) = 1$ and of failure when $c(i, E) = -1$. This function is reinforced at each decision cycle, based on the results of

¹<http://e-ernest.blogspot.fr/2012/03/ernest-111.html>

enaction. The parameters of c constitutes the *Signatures* of interactions. We propose an implementation based on a single layer neuronal network, with a neuron per interaction (1). To this end, the context E_t is coded as a binary vector $[\epsilon_{1,t}; \dots; \epsilon_{n,t}]$, with n the number of interactions and sequences used by the agent, and $\epsilon_{k,t} = 1$ when $i_k \in E_t$. Each signature of interaction or sequence i is composed of a set W_i of n weights $w_{i,k}$ and a bias $w_{i,n+1}$.

$$c(i, E_t) = g \left(\left(\sum_{k \in [1;n]} \epsilon_{k,t} \cdot w_{i,k} \right) + w_{i,n+1} \right) \quad \text{with} \quad g(x) = \frac{2}{1 + e^{-x}} - 1 \quad (1)$$

A set W_i is updated each time the interaction or sequence i is completed as a success or a failure, using the delta rule (2). We note $r_{i,t} = 1$ if i succeeded and $r_{i,t} = -1$ if i failed.

$$w_{a,k}^t \leftarrow w_{a,k}^{t-1} + \alpha \times \epsilon_{k,t-1} \times (r_{a,t} - c(a, E_{t-1})) , \quad \forall k \in [1; n+1], \quad \alpha \in [0; 1] \quad \text{the learning rate.} \quad (2)$$

We propose to use a hard-coded structure, called *Space Memory*[2], that integrates and tracks detected objects, to study more precisely the modified mechanism of signature of interaction and the exploitation mechanism of the SMS. A space memory M is composed of a set P of position p in egocentric reference. A part of these positions is related to positions that correspond to interactions. This subset of positions constitutes the area of space the agent can observe through interactions, we call *Observable Space* P_O . Any geometrical transformation can be approximated by a function $\tau : p \rightarrow p'$. Moving an element in space thus consists in changing its position according to τ . Each primary interaction i produces a transformation τ_i . Note that [1] proposes mechanisms to construct such a structure without any preconception.

The selection mechanism is based on two decisional systems. The exploration system leads the agent to try interactions for the sake of learning signatures. As the frequency of the enaction of sequence of interaction is very low, we propose a mechanism that foster the less tested weights. For each weight of each signature, we attribute a counter $c_{i,k}$ that is incremented each time the weight is modified. At each decision cycle t , we compute the relevance of testing an interaction or a sequence in the current context E_t , defined by $\sum_{k \in [1;n]} c_{i,k} \times \epsilon_{k,t}$. An interaction or a sequence is candidate when its certitude of success is smaller (in absolute value) than a threshold, called *reliability threshold*. The candidate interaction with the lowest relevance is then selected and enacted. If no interaction is candidate, the agent uses the exploitation mechanism.

The exploitation mechanism selects interactions to maximize valence in the short and medium terms. It considers the relative movement of objects generated by the enaction of interactions, and adds a utility value to valence of interactions, that depends on the variation of distance of objects and the valence of interactions afforded by these objects. Closest objects have a greater influence as the agent is more likely to interact with them in the short term. Formally, we note M^τ an *image* of the Spatial Memory M when a transformation $\tau \in T$ is applied, and E^τ the list of interactions stored in M^τ , limited to P_O . For each transformation τ , a distal object that affords an interaction i is considered as present with a certitude of $c(i, E^\tau)$. We note $d(\tau)$ the distance of the object. The utility value Δ_{i_k} , that characterizes the global variation of distance produced by a candidate interaction i_k , is computed as follows:

$$\Delta_{i_k} = \sum_{i \in I} \left(\int_{\tau \in T} c(i, E^{\tau + \tau_{i_k}}) \times f(d(\tau + \tau_{i_k})) - \int_{\tau \in T} c(i, E^\tau) \times f(d(\tau)) \right) \times v_i \quad (3)$$

Where $f : \mathbf{R}^+ \rightarrow]0; 1]$ is a function that characterizes object influence according to their relative distance. In our implementations, we use the function $f : x \rightarrow \gamma^x$ with $\gamma \in]0; 1]$. The mechanism then selects, among candidates i_k , the next intended interaction i_{t+1} defined by $i_{t+1} = \max_{i_k} (v_{i_k} + \beta \times \Delta_{i_k})$, with $\beta \in \mathbf{R}^+$ the *influence coefficient* of the SMS.

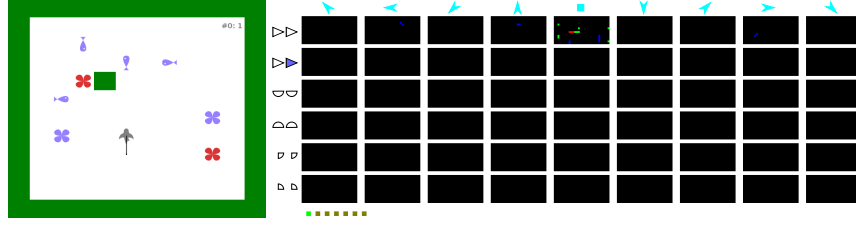


Figure 2: Left: the agent in its environment. The agent is represented as a shark, preys as blue fishes and food as blue alga. Right: the interactional context. Inputs are gathered according to properties of visual sequences of interaction (vertically: related primary sequence, horizontally: observed movement), and each group is topographically organized. Input related to primary sequences are represented with seven squares (here, *move forward twice* is enacted).

4 Implementation on an artificial agent

We have implemented our mechanisms on an agent moving in a 2-dimensional continuous environment (Figure 2). The agent-environment system defines a set of 7 primary interactions (valence is given in parenthesis): move forward of half the length of the agent \triangleright (1), bump \blacktriangleright (-5), move forward and eat something edible \blacktriangleright (50), turn left of 45° \triangleleft (-1), turn right of 45° \triangleleft (-1), turn left of 22.5° \triangleright (-1), turn right of 22.5° \triangleright (-1). We add a set of visual secondary interactions that consists in seeing an element of a color among {red, green, blue} at a position p (supposed defined by a movement and optic flow), while enacting one of the primary interactions (except for *bump* that does not produce movement). We discretize observable positions as a regular grid of 30×15 positions to simplify movements in space memory.

We define 4 types of elements, characterized by a color that makes them recognizable according to the interactional system of the agent: wall block (green) that affords *bump*, food and mobile preys (blue) that affords *eat*, and alga (red) that has no influence on enaction of interactions. Mobile preys move forward at a same speed than the agent. Edible elements are removed when eaten and a same element is added in a randomly selected empty place.

We introduce sequences of interactions to detect relative movements of elements. We only consider relevant sequences. Note that defining relevant sequences is an open question that we intend to address in future works. Primary interactions defines the following sequences: move forward twice $\triangleright\triangleright$ (2), bump \blacktriangleright (-5), move forward, then eat $\triangleright\blacktriangleright$ (52), turn left of 90° $\triangleleft\triangleleft$ (-2), turn right of 90° $\triangleleft\triangleleft$ (-2), turn left of 45° $\triangleright\triangleright$ (-2), and turn right of 45° $\triangleright\triangleright$ (-2). A sequence of secondary interactions is composed of a sequence of two secondary interaction. Blue elements can be immobile or move in 8 directions, relative to the agent, as it can rotate of 45° . The agent can thus experiment, for each position and primary interaction, 9 sequences of secondary interactions. Green and red elements are immobile: the agent can experiment one sequence. We thus define a set of $(9 + 1 + 1) \times 30 \times 15 \times 6 = 29700$ sequences of secondary interactions.

We first test the signature mechanism. We let the agent interacts with its environment, and stop the experiment after 100 000 decision cycles. Figure 3 shows examples of signatures. We can observe that *move forward twice* is related to the absence of green and blue element in front of the agent (mid-red blob on signature), and *bump* is related to a green element in front of the agent. We can note that the size of these elements are nearly the size of the agent. The signature of *move forward then eat* designates a blue element for which the position depends of its movement (and *move forward twice* is related to their absence): the signature thus integrates dynamic properties of preys. *Turn* sequences, that cannot fail, are strongly

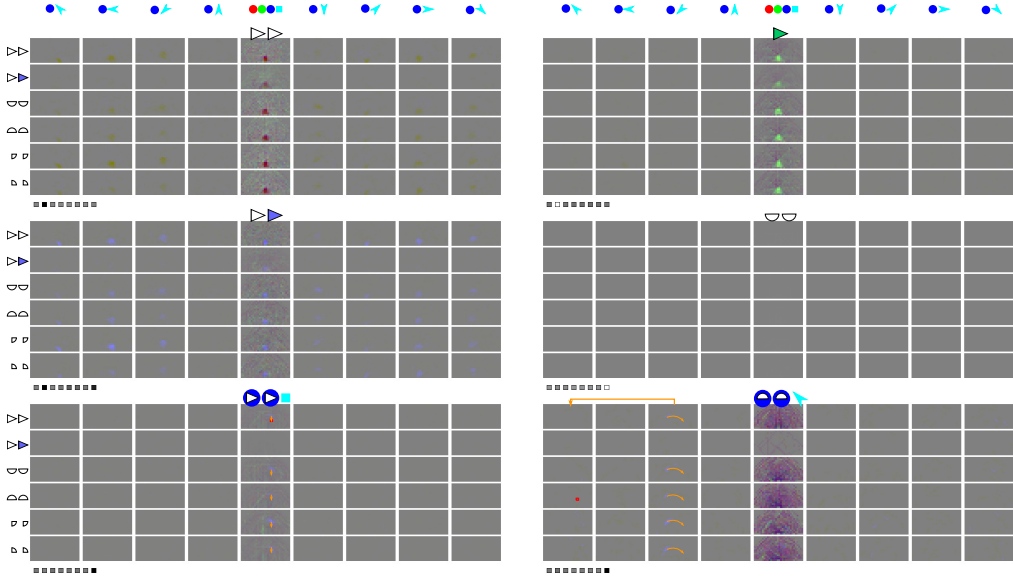


Figure 3: Signatures of sequences of interactions *move forward twice*, *bump*, *move forward then eat*, *turn left of 45°*, *move forward twice* and seeing a blue immobile element, and *turning left of 45°* and seeing a blue element with a forward-left movement (at a position designated by a red square), obtained after 100 000 decision cycles. We gather weights related to visual sequences characterized by a same associated sequence (left) and a same color and observed movement (top). Each group is organized to match the topography of the visual field. Weights related to primary interactions and the bias are represented with squares. A black color means a weight of -1 and a white color means a weight of 1. We can observe that *move forward* is related to the absence of blue and green element in front of the agent, *bump* is afforded by a green element and *eat* by a blue element for which the position depends on the movement of preys. *Turns* sequences are strongly related to the bias as they cannot fail. Signatures of visual sequences designate interactions related to seeing an element of a same color, at a position that correspond to the transformation produced by enacted sequences (orange arrows).

related to the bias. Signatures of visual sequences shows interesting properties: they designate elements of a same color than the considered sequence, but at a position that correspond to the position of the sequence, moved by the movement produced by sequences of interactions and the movement of objects. Figure 4 shows examples of these observed movements. In case of sequences of visual interactions related to *turns* interactions, we also observe a permutation in the movement characterized by sequences designated by signatures. These observations show that signatures integrate spatial and dynamics properties of the environment.

We then add the space memory to test the exploitation mechanism. We propose a simplification of the interactional system: as signatures gather interactions related to a same element, we propose to consider sequences of visual interactions regardless of their associated primary interaction. This simplification decreases the number of sequence and significantly increases the frequency of enaction of these sequences. The downside is that we cannot define signatures of sequences of visual interaction nor enact these sequences. We increase the resolution of the grid that define positions to observe more precisely signatures of interactions.

The experiment is conducted as previously. We let the agent interact with its environment,

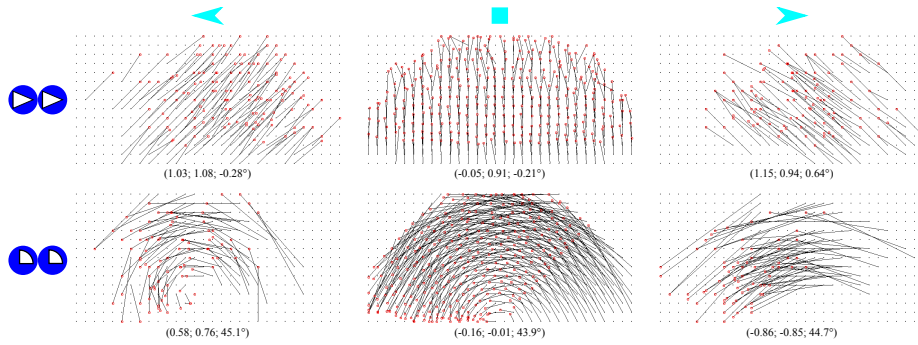


Figure 4: Difference of position between a sequence of visual interaction (black dots) and the barycenter of position of sequences designated by its signature (red circles). Average transformations (x, y, θ) produced by enaction of associated interactions are easily recognizable.

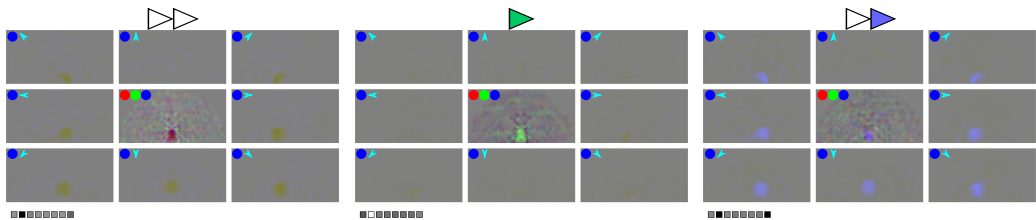


Figure 5: Signatures of sequences *move forward* twice, *bump* and *move forward* then *eat*, obtained after 25 000 decision cycles with the simplified interactional system. Object constructed by the agent are similar to that obtained with the complete interactional system.

driven by a curiosity and an exploitation mechanism. We had to reduce the speed of prey to 0.7 times the speed of the agent to allow the agent to reach a prey by approaching it laterally with an angle of 45° . Figure 5 shows signatures of primary interactions obtained after 25 000 decision cycles. The objects observed by the agent are similar to previously.

We then disable the agent’s learning mechanism and observe its behavior in presence of a prey. It appears that the agent “aims” at a point in space in front of a prey (Figure 6). Indeed, the signature of *eat* interaction considers the movement of preys: the “center” of the object (i.e. the point of space where the interaction can be afforded) is not in front of the prey, but in front of the position of the prey after enacting the interaction. This property is especially visible when we place a wall block in front of the agent (Figure 6 right): as the prey moves left behind the wall block, the agent turns left to reach the prey, although the prey is still in the left side of the agent. This observation shows that the agent considers the future positions of objects rather than their current position.

5 Conclusion

We proposed an adaptation of mechanisms proposed in [2] for a dynamic environment, inspired by sequential RI models. We observed that only a few modifications were needed to allow an artificial agent to discover dynamic properties of its environment. The use of sequences of interactions has a little incidence on principles of mechanisms developed for a static environment.

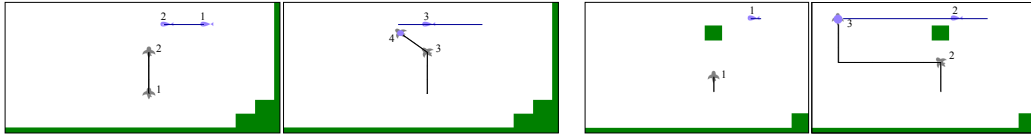


Figure 6: Left: the agent is strongly attracted by the prey (1-2), as it affords an interaction with a high valence. The agent is attracted by the future position of the prey: it turns left while the prey is at its right side (2-3), which allows the agent to reach the prey (4). Right: a wall block is added in front of the agent (1). The agent “anticipates” the position of the prey and bypass the wall by its left side (2). Without perceiving the prey anymore, the agent turns left and moves forward to pass ahead the prey and reach it (3).

Indeed, a sequence of interaction, like an interaction, can success or fail, which makes it possible to use them as an input of the interactional context to complete or replace interactions without modifying the principle of signatures of interactions. The agent shows that it can learn and integrate elements that afford its interactions and spatial properties of its environment, like with mechanisms developed for static environments. It also learns dynamic regularities of its environment, under the form of sequences of interactions that allow to detect movements of objects and to predict their position. The use of a hard-coded space memory is a strong pre-conception. However, we observed that the modifications only affect the interactional system of the agent. The principles of signatures of interactions, of the space memory and of exploitation mechanisms do not need modifications, and stay similar to mechanisms of the agnostic space memory developed previously[1]. In future works, we intend to implement our mechanisms in more complex systems, and in particular agents using continuous sets of interaction.

References

- [1] Simon L. Gay. *Mécanismes d'apprentissage développemental et intrinséquement motivés en intelligence artificielle : étude des mécanismes d'intégration de l'espace environnemental*. PhD thesis, Université Claude Bernard Lyon 1, France, 2014.
- [2] Simon L. Gay, Olivier L. Georgeon, and Christian Wolf. Autonomous object modeling based on affordances for spatial organization of behavior. *IEEE Fourth Joint International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EPIROB)*, 2014.
- [3] Olivier L. Georgeon and D. Aha. The radical interactionism conceptual commitment. *Journal of Artificial General Intelligence*, 2(4):31–36, 2013.
- [4] James J. Gibson. The theory of affordances. In Eds. Robert Shaw and John Bransford, editors, *In Perceiving, Acting and Knowing: Toward an Ecological Psychology*, 1977.
- [5] Michael S. A. Graziano, Charlotte S. R. Taylor, and Tirin Moore. Complex movements evoked by microstimulation of precentral cortex. *Neuron*, 34(5):841–851, 2002.
- [6] David P. M. Northmore. The optic tectum. In Academic Press, editor, *In : Encyclopedia of Fish Physiology: from Genome to Environment*, 2011.
- [7] Pierre-Yves Oudeyer, Frdric Kaplan, and Verena V. Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(2):265–286, 2007.
- [8] Jean Piaget. *The construction of reality in the child*. Basic Books, New York, 1954.
- [9] Emre Uğur, Mehmet R. Doğar, Maya Çakmak, and Erol Şahin. The learning and use of traversability affordance using range images on a mobile robot. *IEEE International Conference on Robotics and Automation*, pages 1721–1726, 2007.